Нормы платформы

Мы понимаем свою роль в том, чтобы предотвращать распространение языка вражды. Вместе с небезразличными пользователями наша команда беспрерывно следит, чтобы на платформе не было места травле, прославлению насилия и призывам к ненависти.

Как мы модерируем контент

Если вы столкнулись с угрозами или оскорблениями, подверглись нападкам или увидели другие вредоносные материалы, сообщите нам об этом **с помощью кнопки «Пожаловаться»**. Она есть у профилей, сообществ, записей, комментариев, фотографий, видео и историй. Мы проверяем каждое обращение: зачастую за несколько минут, обычно — в пределах часа. Если после жалобы мы заблокируем профиль или сообщество, то сообщим об этом автору обращения.

Мы разделяем позицию наших пользователей, клиентов и партнёров касательно борьбы с распространением языка вражды. Чтобы усилить модерацию, мы добавили отдельную причину для жалоб на такие материалы — **«Враждебные высказывания»**.

Наша команда не только рассматривает поступающие жалобы, но и **самостоятельно отслеживает появление вредоносного контента**. Делать это круглосуточно нам помогают механизмы автоматического поиска, технология цифровых отпечатков и нейронные сети.

Экспериментальная функция для борьбы с оскорбительными комментариями позволила сократить их количество на 36% и дополнительно исследовать

материалов, которые провоцируют комментаторов использовать такую риторику. Новая технология призвана **оградить пользователей** от разжигающего ненависть контента; **обезопасить рекламодателей и партнёров** от размещения объявлений в сообществах, где проявляется нетерпимость; ускорить автоматическую блокировку запрещённых материалов и повысить эффективность модерации.

Какие материалы запрещены на платформе

ВКонтакте — глобальная платформа, которой пользуются люди из разных стран, с разными взглядами, убеждениями и мнениями. Мы понимаем, что у каждого — свои представления о хорошем и плохом, о нормах морали в целом. Чтобы поддерживать безопасную среду для всех, у нас действуют **единые правила**.

ВКонтакте **запрещено распространять угрозы применения насилия и враждебные высказывания**, а также преследовать или унижать человека либо группу людей — в том числе на основании религиозной, культурной, расовой, этнической или национальной принадлежности, сексуальной и гендерной идентичности, наличия особенностей физического развития или заболеваний.

Пользуясь ВКонтакте, **воздержитесь** также от публикации материалов, которые содержат:

- угрозы совершения кражи, вандализма, причинения физического вреда или финансового ущерба;
- поощрение самоубийств или нанесения увечий себе, другим людям либо животным, а также изображения подобных увечий и другой шокирующий контент;
- личную информацию других людей, размещённую с целью шантажа или преследований, — в том числе угрозы публикации интимных изображений;
- пропаганду незаконной деятельности и преступных организаций, в том числе терроризма;
- информацию об употреблении или продаже наркотических средств, предложения купить или продать оружие, боеприпасы или взрывчатые вещества;
- спам, мошенничество или заведомо ложную информацию;

- нарушения авторских прав или прав на товарные знаки;
- другие нарушения, предусмотренные Правилами ВКонтакте или нормами законодательства.

Чтобы не столкнуться с ограничением доступа к профилю, воздержитесь от публикации вредоносного и оскорбительного контента, проявлений расизма, гомофобии, сексизма и других враждебных высказываний. Обратите внимание, что размещение некоторых материалов не только мешает людям спокойно общаться, но и может повлечь юридическую ответственность. Подробнее о таком контенте — в пункте 6.3.4 Правил ВКонтакте. Материалы подобного рода **недопустимы и в качестве рекламных объявлений**.

Противодействие разжиганию вражды и ненависти

Профили и сообщества ВКонтакте — **не место для травли и прославления насилия**. Мы запрещаем формировать на нашей платформе объединения и продвигать организации, которые распространяют идеологию вражды, насилия, преследования людей и вмешательства в чужую частную жизнь. Даже в виде закрытых групп.

Когда мы рассматриваем жалобы на материалы, профили и сообщества, то **учитываем контекст**. Например, если публикация осуждает травлю, расизм, гомофобию или сексизм либо привлекает внимание к этим проблемам, она может оставаться публично доступной. Кроме того, мы с особым вниманием относимся к образовательному контенту, художественным произведениям, историческим и общественно-резонансным материалам. Важно, чтобы они были созданы для отображения и обсуждения фактов и происшествий, а не преследования людей и ущемления чьих-то прав.

При этом мы стараемся **предостерегать** пользователей от материалов, которые могут шокировать или показаться неприемлемыми, даже если они размещены в информационных или образовательных целях и не призваны кого-либо оскорбить. Для этого мы тестируем систему предупреждений. К примеру, включая видео, пользователь может встретить уведомление, что в материале есть сцены насилия или другой тяжёлый контент. Такое предупреждение может появиться, допустим, перед роликом с кадрами операции из закрытого сообщества студентов

распространяют идеологию ненависти и вражды. Если пользователь грубо или неоднократно нарушает нормы платформы, его профиль или сообщество **будут заблокированы** навсегда, без права на восстановление доступа.

Заблокировано по теме разжигания ненависти и вражды:

292 000 единиц контента | **680** личных страниц | **1 260** сообществ

Первый квартал 2020 года

228 500 единиц контента | **660** личных страниц | **1 210** сообществ

Второй квартал 2020 года

Как проявляется риторика насилия и ненависти

Язык вражды может выглядеть по-разному — и мы всегда учитываем контекст. Наши модераторы смотрят, есть ли в материалах признаки, которые подтверждают враждебность намерений автора высказывания:

- неприязнь, основанная на каких-либо особенностях или различиях;
- оскорбительное поведение, презрение к чужим ценностям и взглядам;
- выражение собственного превосходства, которое сопровождается необоснованным и несправедливым отношением к конкретному человеку или группе людей.

Мы **блокируем** профили и сообщества, через которые распространяют:

- прямые угрозы жизни и здоровью (в частности, они могут проявляться обещанием «уничтожить»);
- призывы к суициду в различной форме (предложение или настояние что-то сделать с собой, а также описание способов самоубийства);
- враждебные, угрожающие и поощряющие насилие высказывания, нападки на человека или группу людей с целью унижения человеческого достоинства или утверждения о неполноценности (такие материалы могут содержать ссылку на профиль, фотографии человека, его адрес проживания или номер телефона,

- призывы к изоляции или сегрегации (к примеру, чтобы определенные люди жили отдельно: «вывезти их», «пусть там и живут у себя в… и не высовываются» и прочее), пожелания серьёзного вреда и призывы нанести его, поощрение преследования или оскорбительного поведения, скрытые призывы или подстрекательства к насилию (зачастую сопровождаются призывами «гнать» таких людей, «отправлять на свалку» и т. п.);
- словесные подтверждения превосходства одних групп над другими для оправдания насилия, дискриминации, сегрегации или изоляции на основании религиозной, этнической или национальной принадлежности, сексуальной и гендерной идентичности, наличия особенностей физического развития либо заболеваний (они могут проявляться в виде сравнения конкретной группы людей с насекомыми, грязью, нелюдями, низшим сортом и подобной риторики).

Такие высказывания недопустимы на нашей платформе, в том числе в виде шуток и мемов. Иногда важно, чтобы человек, который потенциально мог стать объектом травли, сам пожаловался на материал. Это помогает определить наличие враждебных намерений и становится обоснованием для удаления публикации.

Ваше неравнодушие и наши инструменты делают ВКонтакте безопасным и комфортным местом. Если вы столкнулись с оскорблениями, сообщите об этом с помощью кнопки «Пожаловаться». Если же вам или кому-то другому прямо сейчас грозит реальная опасность, не теряйте времени и сообщите об этом в экстренные службы.

Здоровая среда